

Physical Mapping



Eric D. Green, M.D., Ph.D.
Chief, Genome Technology Branch (NHGRI)
Director, NIH Intramural Sequencing Center (NISC)
Phone: 301-402-0201
FAX: 301-402-4735
E-Mail: egreen@nhgri.nih.gov

Physical Mapping: Outline

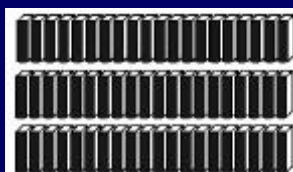
- I. Fundamentals of Physical Mapping
- II. Radiation Hybrid (RH) Mapping
- III. Clone-based Physical Mapping
 - A. Cloning Systems
 - B. Strategies for Clone-based Physical Mapping
 - C. Clone-based Physical Maps of Mammalian Genomes
- IV. Future Prospects

Physical Mapping: Goals

- Stress the Practical Aspects of Physical Mapping
- Focus More on the Mapping of Mammalian Genomes
- Highlight Relevant Literature
- Provide Information on Relevant Electronic Resources

Genome Sizes

Human Genome
Mouse Genome



~3,000,000,000 bp

Fruit Fly Genome



~160,000,000 bp

Nematode Genome



~100,000,000 bp

Yeast Genome

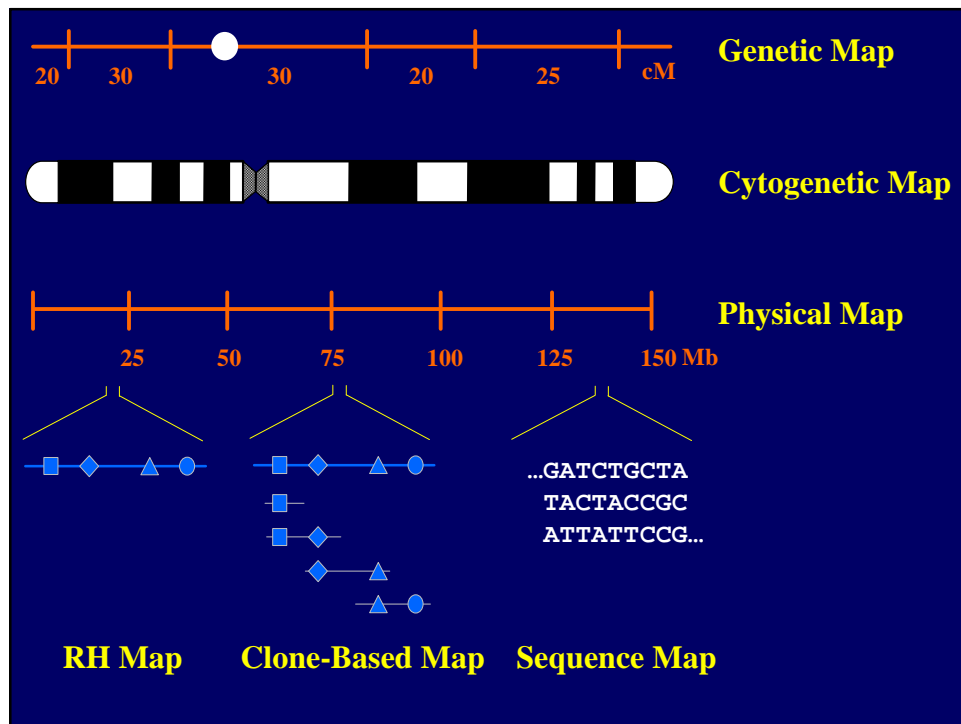


~15,000,000 bp

***E. coli* Genome**



~5,000,000 bp



Fundamentals of Physical Mapping

- Importance of Physical Maps:
 - Localization and Isolation of Genes (e.g., Positional Cloning)
 - Study of Genome Organization and Evolution
 - Framework for Systematic DNA Sequencing
- Mapping is About Order
- Physical Mapping Involves:
 - Ordering of Clones and/or Landmarks
 - Typically with Some Physically Measurable Metric
- General Types of Physical Maps:
 - Landmark Only
 - Clone-based
 - Sequence

Landmark Only Physical Maps

- Restriction Mapping by Pulsed-Field Gel Electrophoresis

Riethman et al. (1997) *Genome Analysis*, Vol. 1, Chap. 2

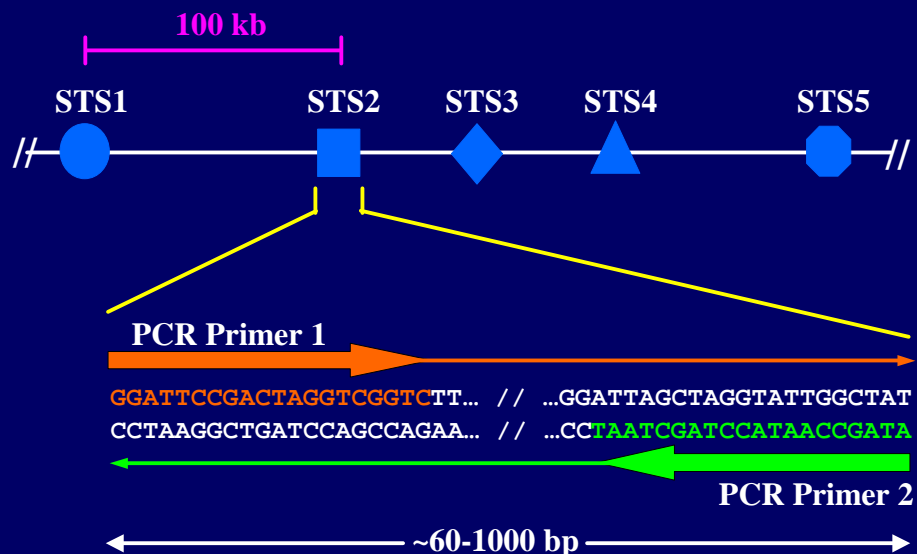
- Radiation Hybrid (RH) Mapping

Matise et al. (1999) *Genome Analysis*, Vol. 4, Chap. 6

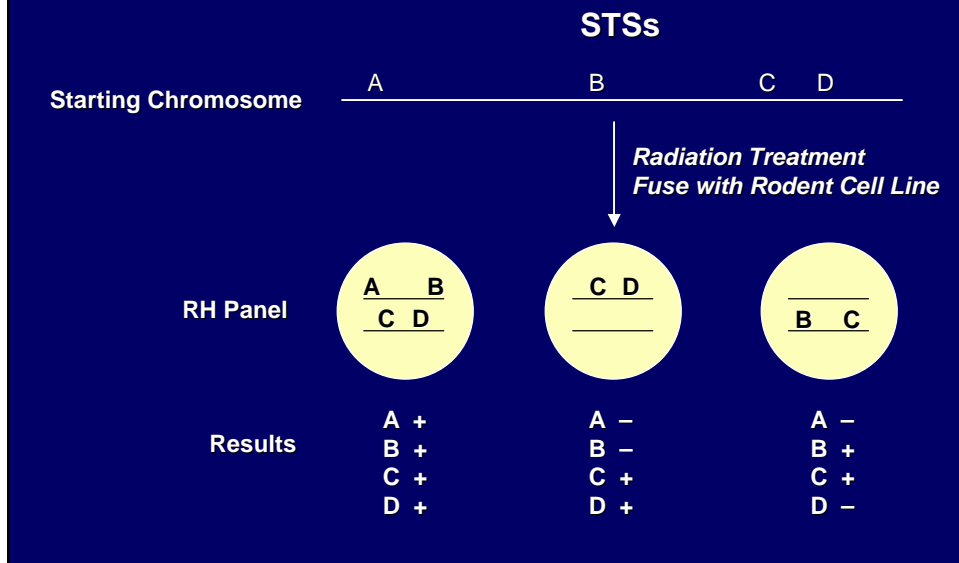
Lecture by Dr. Tara Matise, *Current Topics '99*
(see www.nhgri.nih.gov/COURSE99)

Stanford University Genome Center
(see shgc-www.stanford.edu)

Sequence-Tagged Sites (STSs)



RH Mapping: Theory



‘Classic’ Human RH Mapping Panels

	Genebridge 4 (GB4)	Stanford G3 (G3)
X-ray dosage	3,000 rad	10,000 rad
Map units	3,000 cR	10,000 cR
Cell lines	93	83
Average retention	32%	16%
Average fragment size	25 Mb	2.4 Mb
Effective resolution	1 Mb	0.25 Mb
<i>Utility</i>	<i>Long-range continuity</i>	<i>Higher resolution</i>

RH Mapping: Available Resources

- DNA from RH Panels

Research Genetics: www.resgen.com

- RH Mapping Servers

Human:

www.sanger.ac.uk/Software/RHserver/RHserver.shtml

www-shgc.stanford.edu/RH/G3index.html

carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl

Mouse:

www.genome.wi.mit.edu/cgi-bin/mouse_rh/rhmapauto/rhmapper.cgi

Rat:

rgd.mcw.edu/RHMAPSERVER/

- General Reference Information: compgen.rutgers.edu/rhmap

RH Mapping-based Human Gene Map



A Physical Map of 30,000 Human Genes

P. Deloukas,^a G. D. Schuler, G. Gyapay, E. M. Boerley, C. Soderlund, P. Rodriguez-Tome, L. Hui, T. C. Matise, K. B. McKusick, J. S. Beckmann, S. Bontalio, M.-T. Bihoreau, B. B. Birren, J. Brown, A. Butler, A. B. Castle, N. Chiannelli, C. Clee, P. J. R. Day, A. Dehejia, T. Dibbing, N. Drouot, S. Duprat, C. Flaxman, S. Fox, S. Gelling, L. Green, P. Harrison, R. Hocking, E. Holloway, S. Hunt, S. Kell, P. Lijnzaad, C. Louis-Dit-Sully, J. Ma, A. Mendis, J. Miller, J. Morisette, D. Muscat, H. C. Nussbaum, A. Peck, S. Rosen, D. Simon, D. R. Skovlin, R. Staples, L. D. Stein, E. A. Stewart, M. A. Suchard, T. Thangarajah, N. Vega-Carrizosa, C. Webber, X. Wu, J. Hudson, C. Aulfray, N. Navara, J. M. Sibel, M. H. Polymeropoulos, M. R. James, E. S. Lander, T. J. Hudson, R. M. Myers, D. R. Cox, J. Weissenbach, M. S. Boguski, D. R. Bentley

Science 282:744-746, 1998

www.ncbi.nlm.nih.gov/genemap99

Radiation hybrid map of the mouse genome

William J. Van Etten¹, Robert G. Steen¹, Huy Nguyen¹, Andrew B. Castle¹, Donna K. Slemon¹, Bing Ge², Chad Nusbaum¹, Greg D. Schoket¹, Eric S. Lander^{1,2} & Thomas J. Hudson^{1,2}

Nature Genetics
22:384-387 (1999)

A radiation hybrid map of the rat genome containing 5,255 markers

Takashi K. Watanabe^{1*}, Marie-Thérèse Bihoreau^{2*}, Linda C. McCarthy^{3,4*}, Susanna E. Rognes^{2*}, George H. Hwang¹, Anusht Tajji¹, Julie Brown², Yuki Yanagaki¹, Ayako Moriguchi-Miyakita¹, Koiko Ogi¹, Yoshiko Ooi¹, Shiro Okumura¹, Naohide Kasamatsu¹, Ei-ichi Takahashi¹, Kazuo Taniuchi¹, Hiroshi Hagiwara¹, Masakazu Adachi¹, Cathy Wittber², Maria Davis², Suzanne Ruel², Catherine Krugler², Angela Smith¹, Ricky Crisler¹, Jonathan Miller², Thiru Thangarajah², Philip J.R. Day², James R. Hansen^{1,4}, Yoonsoo Lee¹, Eishi Taki², Yusuke Nakamura⁵, Peter N. Goodfellow⁶, G. Mark Lathrop⁷, Akira Taniguchi¹ & Michael R. James²

Nature Genetics
22:27-36 (1999)

A High-Density Integrated Genetic Linkage and Radiation Hybrid Map of the Laboratory Rat

Robert G. Steen,^{1,*} Anne E. Kwitek-Black,^{2,*} Christopher Glenn,^{3,*} Jo Gullings-Handley,² William Van Etten,¹ Q. Scott Atkinson,² Diane Appel,¹ Simon Twigger,² Melanie Muir,¹ Tim Mull,² Mary Gamados,² Mushira Kisebah,² Keni Russo,² Robbin Crane,¹ Michael Popp,² Marc Peden,² Tara Mattise,⁴ Donna M. Brown,³ Jian Lu,² Stephen Kingsmore,¹ Peter J. Tonellato,² Steve Rozen,¹ Donna Slonim,¹ Peter Young,¹ Margit Knoblauch,² Abraham Provoost,⁷ Detlev Ganten,⁶ Steven D. Colman,¹ Jonathan Rothberg,² Eric S. Lander,¹ and Howard J. Jacob,^{2,9}

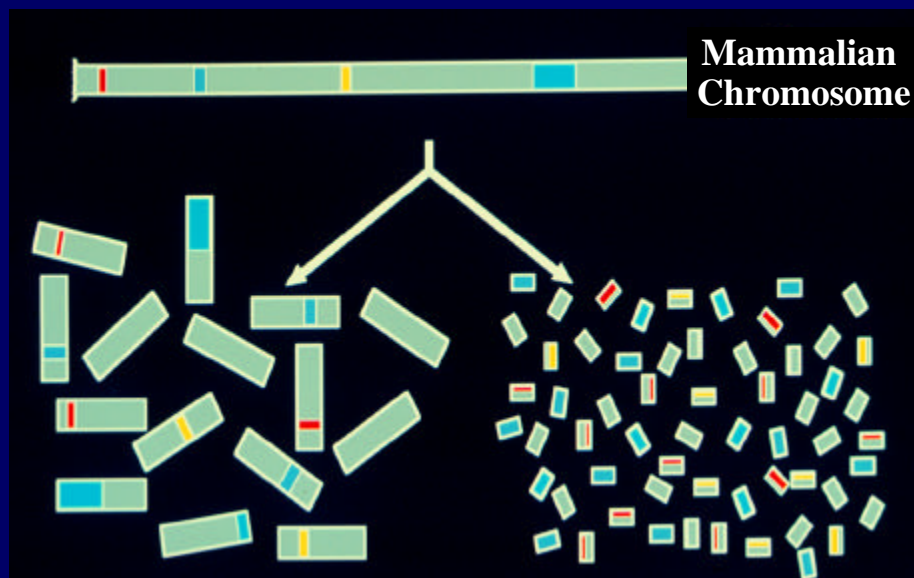
Genome Research
9:AP1-AP8 (1999)

A radiation hybrid map of the zebrafish genome

Jablon Geller¹, Gert-Jing Ruedi¹, Henwig Beaz¹, Fouke von Böhmer¹, Linda Brod¹, Marcus P.S. Dobens², Karle Finger¹, Corinne Hilde¹, Michael A. Gies¹, Horn Geiger¹, Silke Geiger-Rudolph¹, Doreen Gildou¹, Stefanie Glaser¹, Lara Götting¹, Hiltrich Habbeck¹, Katy Hänge¹, Scott Holley¹, Jeremy Kenna¹, Aasta Kim¹, Folger Knorr¹, David Lohbauer¹, Florian Maderbacher¹, Ulrike Marzys¹, Sophia Neuhaus¹, Carl Neumann¹, Jozsef Nicolau¹, Francisco Poligr¹, Russell Ray¹, Jero M. Rick¹, Thany Reich¹, Tobias Rocco¹, Heide E. Schauerle¹, Alexander T. Schier¹, Ulrike Schöneberg¹, Hella-Timm-Schneiders¹, Stefan Schulte-Morke¹, Catrin Seyler¹, William S. Talbot¹, Christian Waller¹, Christine Kretzschmar¹ & Pascal Haller¹

Nature Genetics
23:86-89 (1999)

Jigsaw Puzzle Analogy of Clone-based Physical Mapping



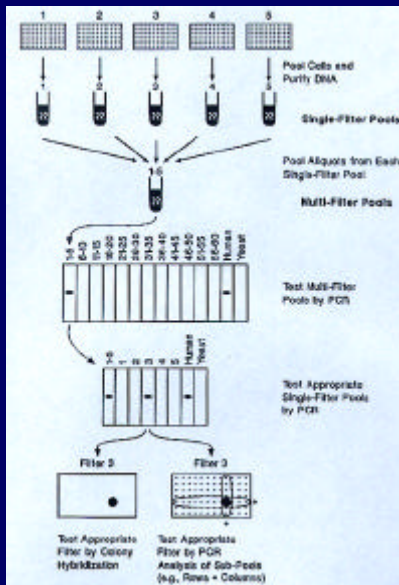
Clones for Physical Mapping: General Points

- **Want Cloned DNA to Accurately Reflect the Starting Genome**
 - Problem of Instability**
 - Problem of Chimerism**
- **Development of ‘Array Mentality’ for Clone Libraries**
 - Clones Arrayed in Individual Wells of Microtiter Plates**
 - Various Densities (e.g., 96-and 384-Well Plates)**
- **Advantages of Arrayed Libraries (‘Reference Libraries’)**
 - 1. Simplicity of Storing and Transferring Clone Collections**
 - 2. Convenient Format for Retrieving Clones of Interest**
 - 3. Ability to Assimilate Data on Common Clones**
 - 4. Repeated PCR-based Screening**
 - 5. Repeated Hybridization-based Screening**

Commercial Involvement in Clone Distribution

Research Genetics:	www.resgen.com
Incyte Genomics:	www.incyte.com
ATCC:	www.atcc.org
BACPAC Resource:	www.chori.org/bacpac

Screening Clone Libraries: PCR-based Approaches



Cosmids

- Bacterial-based Cloning System
- ‘Antique’ of the Large DNA Cloning Systems
- Plasmid Vector with Bacteriophage Packaging Sequences (*cos* Sites)
- High-Efficiency Packaging System
 - Relatively Homogeneous Insert Sizes
 - Libraries from Small Amounts of DNA (e.g., Flow-Sorted DNA)
 - Antibiotic Selection
- Cloned Inserts: 35-45 kb, Circular DNA
- High Copy Number
 - High Yields of DNA by Standard Methods
 - Instability Problems (Despite Recombination-Deficient Hosts)
- Relatively Non-Chimeric
- Various Libraries (Whole Genomes, Individual Chromosomes)
- References: Sambrook et al. (1989), Wahl et al. (1987), Ivens et al. (1993), Evans (1998)
- ‘Fosmids’ [Kim et al. (1992)]: Cosmid Vector Engineered with F Factor [Low Copy → More Stable]

P1 Clones

- Bacterial-based Cloning System
- Developed by Sternberg (1990)
- P1-based Vector and Complex P1 Packaging Extracts
 - Limited to 100 kb (Constraints of Viral Particle)
 - 2 *loxP* Sites Results in Circularization of DNA
 - Antibiotic Selection
- Cloned Inserts: 70-100 kb, Circular DNA
- Low Copy Number
 - Low Yields of DNA by Standard Methods
 - Highly Stable (with Recombination-Deficient Hosts)
 - Potential for IPTG Induction → 10-30 Fold Increase
- Relatively Non-Chimeric
- Human and Mouse Libraries Commercially Available
- References: Sternberg (1990), Sternberg et al. (1990), Shepherd et al. (1994), Sternberg (1998)

P1-Derived Artificial Chromosomes (PACs)

- Bacterial-based Cloning System
- Developed by Ioannou et al. (1994)
- Slightly Modified P1 Vector
 - Lacks Packaging Signal
 - Antibiotic Selection
- Transform by Electroporation
- No Packaging of DNA → Larger Size Capacity
- Cloned Inserts: 100-150 kb, Circular DNA
- Low Copy Number
 - Low Yields of DNA by Standard Methods
 - Highly Stable (with Recombination-Deficient Hosts)
- Relatively Non-Chimeric

Bacterial Artificial Chromosomes (BACs)

- Bacterial-based Cloning System
- Developed by Shizuya et al. (1992)
- Based on the *E. coli* F Factor (Fertility Plasmid): Replication Control
- BAC Vectors
 - Cloning site in LacZ Gene (Blue/White Selection)
 - Antibiotic Selection
- Transform by Electroporation
- No Packaging of DNA → Larger Size Capacity
- Cloned Inserts: 100-200 kb, Circular DNA
- Low Copy Number
 - Low Yields of DNA by Standard Methods
 - Highly Stable (with Recombination-Deficient Hosts)
- Relatively Non-Chimeric
- Numerous Libraries Available (see www.chori.org/bacpac)
- See Birren et al. (1998)

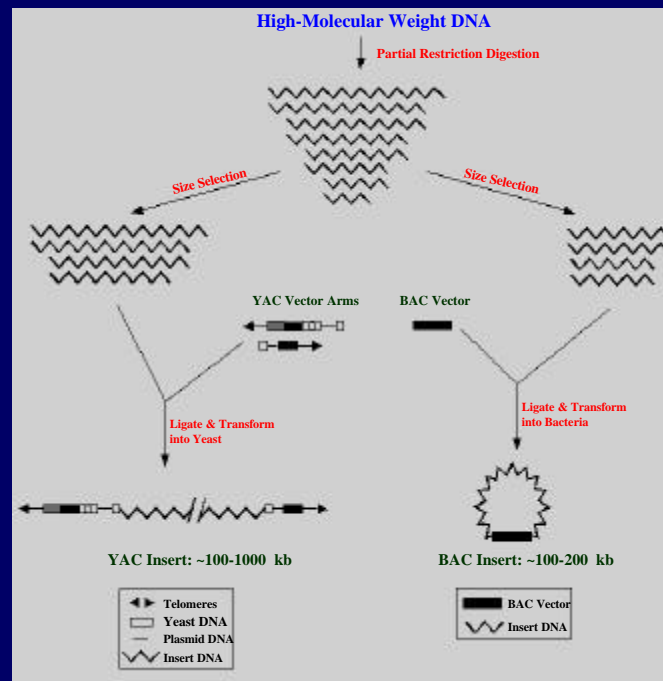
Yeast Artificial Chromosomes (YACs)

- Yeast-based Cloning System (*Saccharomyces cerevisiae*)
- Developed by Burke et al. (1987)
- System Based on Ability to ‘Harness’ Cloned DNA with Structural Elements Required for the Propagation of a Linear Chromosome in Yeast
- Cloned Insert: ~100 kb to >1,000 kb, Linear DNA
- Spheroplast Transformation Procedure
 - Technically Demanding
 - Poorly Defined Upper Size Limit for Cloned Insert
- References: Hieter et al. (1990), Ramsay & Wicking (1991), Schlessinger & Kere (1992), Green et al. (1998)

Major Features of YACs

- Cloned DNA in Single Copy within Yeast Genome
 - Generally Same Structure and Size as Endogenous Chromosomes
 - Limited ‘Access’ to Cloned DNA (e.g., Gel Isolation)
- Chimerism as Major ‘Problem’ (Green et al., 1991)
 - Upwards of 40-60% of Clones in Total Mammalian DNA Libraries
- Instability (e.g., Internal Deletions) as Minor ‘Problem’
- Various Human, Mouse, Rat, (and Other) Libraries Constructed
 - Human:
 - Washington University [Burke and Olson (1991), Brownstein et al. (1989)]
 - CEPH (Includes ‘Mega-YACs’) [Albertsen et al. (1990), Dausset et al. (1992)]
 - ICRF [Larin et al. (1991)]
 - ICI [Anand et al. (1989), Anand et al. (1990)]
 - Mouse:
 - Princeton [Burke et al. (1991), Rossi et al. (1992)]
 - St. Mary’s [Chartier et al. (1992)]
 - ICRF [Larin et al. (1991, 1993)]
 - Whitehead [Kusumi et al. (1993), Haldi et al. (1996)]
 - Rat:
 - Harvard [Cai et al. (1997)]
 - Whitehead [Haldi et al. (1997, 1997)]

Construction of Large-Insert Clones: YACs and BACs



Genome Sizes

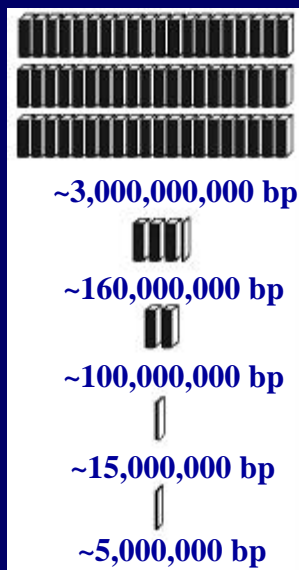
Human
Mouse

Fruit Fly

Nematode

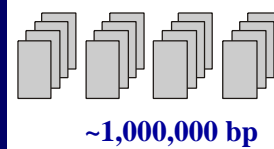
Yeast

E. coli

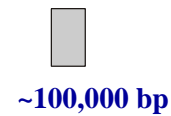


Cloning Capacity

YAC



BAC



Cosmid



Bacteriophage



Strategies for Clone-based Physical Mapping

- Two Key Components ('Jigsaw Puzzle Analogy')
 - Cloned Fragments (Pieces of the Puzzle)
 - Landmarks (Provide Clues for Aligning Pieces)
- Involves the Use of Landmarks to Assembly Clone 'Contigs'
 - Contig*: Overlapping Set of Clones that Together Contains a Contiguous Segment of the Source Genome
- Nature of Landmarks
 - Must Provide 'Unique' Information About the DNA
 - Must be Easy to Identify
- Early Candidates for Landmarks: Restriction Sites
 - E. Coli* [Kohara et al. (1987)]
 - Yeast [Olson et al. (1986), Riles et al. (1993)]
 - Nematode [Coulson et al. (1986)]

Early Physical Mapping of Human Chromosomes

- Strategies Analogous to those Used with *E. coli*, Yeast, and Nematode
 - Applied to Several Human Chromosomes
 - Cosmid Clones (e.g., Flow-Sorted Libraries)
 - Restriction Map Construction and/or Fingerprint Analysis [e.g., Stallings et al. (1990)]
- Shift in Strategies with the Development of YACs
- Distinguishing Features of YACs: No Ability to Readily Purify Cloned DNA
- Modified Fingerprint-based Strategies Attempted with YACs [e.g., Bellanne-Chantelot et al. (1992)]
 1. Requires Gel-Transfer Hybridization
 2. Typically Uses Repetitive Element-Specific Probe(s)
 - Establish YAC 'Fingerprint' → Infer Overlap(s) with Other YACs
- Development of PCR → Sequence-Tagged Sites (STSs)
- 'Common Language' of STSs Proposed by Olson et al. (1989)

(E.D. Green & P. Green, PCR Meth. Applic. 1:77, 1991)

The diagram illustrates the process of YAC mapping. It begins with a **YAC Library** containing various sized clones. A **Screen with PCR Assays for Set of STSs** identifies clones containing specific STSs (STS1, STS2, STS3, STS4). These are then **Assimilate Data** to create a map, and finally **Assemble YAC Contig** to form a continuous sequence of overlapping clones.

The chart displays the following flights and their approximate durations:

- 1968:** Apollo 8 (Dec 16-17), Apollo 9 (Nov 9-10), Apollo 10 (May 18-19), Apollo 11 (Nov 14-16), Apollo 12 (Dec 14-15), Apollo 13 (Apr 11-14), Apollo 14 (Feb 16-17), Apollo 15 (Jul 14-17), Apollo 16 (Apr 12-15), Apollo 17 (Dec 14-17).
- 1970s:** Skylab 2 (May 14-17), Skylab 3 (Jul 28-29), Skylab 4 (Oct 11-12), Apollo 18 (May 14-15), Apollo 19 (May 16-17), Apollo 20 (May 18-19), Apollo 21 (May 20-21), Apollo 22 (May 22-23), Apollo 23 (May 24-25), Apollo 24 (May 26-27), Apollo 25 (May 28-29), Apollo 26 (May 30-31), Apollo 27 (Jun 1-2), Apollo 28 (Jun 3-4), Apollo 29 (Jun 5-6), Apollo 30 (Jun 7-8), Apollo 31 (Jun 9-10), Apollo 32 (Jun 11-12), Apollo 33 (Jun 13-14), Apollo 34 (Jun 15-16), Apollo 35 (Jun 17-18), Apollo 36 (Jun 19-20), Apollo 37 (Jun 21-22), Apollo 38 (Jun 23-24), Apollo 39 (Jun 25-26), Apollo 40 (Jun 27-28), Apollo 41 (Jun 29-30), Apollo 42 (Jul 1-2), Apollo 43 (Jul 3-4), Apollo 44 (Jul 5-6), Apollo 45 (Jul 7-8), Apollo 46 (Jul 9-10), Apollo 47 (Jul 11-12), Apollo 48 (Jul 13-14), Apollo 49 (Jul 15-16), Apollo 50 (Jul 17-18), Apollo 51 (Jul 19-20), Apollo 52 (Jul 21-22), Apollo 53 (Jul 23-24), Apollo 54 (Jul 25-26), Apollo 55 (Jul 27-28), Apollo 56 (Jul 29-30), Apollo 57 (Jul 31-Aug 1), Apollo 58 (Aug 3-4), Apollo 59 (Aug 5-6), Apollo 60 (Aug 7-8), Apollo 61 (Aug 9-10), Apollo 62 (Aug 11-12), Apollo 63 (Aug 13-14), Apollo 64 (Aug 15-16), Apollo 65 (Aug 17-18), Apollo 66 (Aug 19-20), Apollo 67 (Aug 21-22), Apollo 68 (Aug 23-24), Apollo 69 (Aug 25-26), Apollo 70 (Aug 27-28), Apollo 71 (Aug 29-30), Apollo 72 (Sep 1-2), Apollo 73 (Sep 3-4), Apollo 74 (Sep 5-6), Apollo 75 (Sep 7-8), Apollo 76 (Sep 9-10), Apollo 77 (Sep 11-12), Apollo 78 (Sep 13-14), Apollo 79 (Sep 15-16), Apollo 80 (Sep 17-18), Apollo 81 (Sep 19-20), Apollo 82 (Sep 21-22), Apollo 83 (Sep 23-24), Apollo 84 (Sep 25-26), Apollo 85 (Sep 27-28), Apollo 86 (Sep 29-30), Apollo 87 (Oct 1-2), Apollo 88 (Oct 3-4), Apollo 89 (Oct 5-6), Apollo 90 (Oct 7-8), Apollo 91 (Oct 9-10), Apollo 92 (Oct 11-12), Apollo 93 (Oct 13-14), Apollo 94 (Oct 15-16), Apollo 95 (Oct 17-18), Apollo 96 (Oct 19-20), Apollo 97 (Oct 21-22), Apollo 98 (Oct 23-24), Apollo 99 (Oct 25-26), Apollo 100 (Oct 27-28).
- 1980s:** STS-1 (Apr 12-13), STS-2 (Apr 14-15), STS-3 (Apr 16-17), STS-4 (Apr 18-19), STS-5 (Apr 20-21), STS-6 (Apr 22-23), STS-7 (Apr 24-25), STS-8 (Apr 26-27), STS-9 (Apr 28-29), STS-10 (Apr 30-May 1), STS-11 (May 3-4), STS-12 (May 5-6), STS-13 (May 7-8), STS-14 (May 9-10), STS-15 (May 11-12), STS-16 (May 13-14), STS-17 (May 15-16), STS-18 (May 17-18), STS-19 (May 19-20), STS-20 (May 21-22), STS-21 (May 23-24), STS-22 (May 25-26), STS-23 (May 27-28), STS-24 (May 29-30), STS-25 (May 31-Jun 1), STS-26 (Jun 3-4), STS-27 (Jun 5-6), STS-28 (Jun 7-8), STS-29 (Jun 9-10), STS-30 (Jun 11-12), STS-31 (Jun 13-14), STS-32 (Jun 15-16), STS-33 (Jun 17-18), STS-34 (Jun 19-20), STS-35 (Jun 21-22), STS-36 (Jun 23-24), STS-37 (Jun 25-26), STS-38 (Jun 27-28), STS-39 (Jun 29-30), STS-40 (Jul 1-2), STS-41 (Jul 3-4), STS-42 (Jul 5-6), STS-43 (Jul 7-8), STS-44 (Jul 9-10), STS-45 (Jul 11-12), STS-46 (Jul 13-14), STS-47 (Jul 15-16), STS-48 (Jul 17-18), STS-49 (Jul 19-20), STS-50 (Jul 21-22), STS-51 (Jul 23-24), STS-52 (Jul 25-26), STS-53 (Jul 27-28), STS-54 (Jul 29-30), STS-55 (Aug 1-2), STS-56 (Aug 3-4), STS-57 (Aug 5-6), STS-58 (Aug 7-8), STS-59 (Aug 9-10), STS-60 (Aug 11-12), STS-61 (Aug 13-14), STS-62 (Aug 15-16), STS-63 (Aug 17-18), STS-64 (Aug 19-20), STS-65 (Aug 21-22), STS-66 (Aug 23-24), STS-67 (Aug 25-26), STS-68 (Aug 27-28), STS-69 (Aug 29-30), STS-70 (Sep 1-2), STS-71 (Sep 3-4), STS-72 (Sep 5-6), STS-73 (Sep 7-8), STS-74 (Sep 9-10), STS-75 (Sep 11-12), STS-76 (Sep 13-14), STS-77 (Sep 15-16), STS-78 (Sep 17-18), STS-79 (Sep 19-20), STS-80 (Sep 21-22), STS-81 (Sep 23-24), STS-82 (Sep 25-26), STS-83 (Sep 27-28), STS-84 (Sep 29-30), STS-85 (Oct 1-2), STS-86 (Oct 3-4), STS-87 (Oct 5-6), STS-88 (Oct 7-8), STS-89 (Oct 9-10), STS-90 (Oct 11-12), STS-91 (Oct 13-14), STS-92 (Oct 15-16), STS-93 (Oct 17-18), STS-94 (Oct 19-20), STS-95 (Oct 21-22), STS-96 (Oct 23-24), STS-97 (Oct 25-26), STS-98 (Oct 27-28), STS-99 (Oct 29-30), STS-100 (Nov 1-2).
- 1990s:** STS-101 (Nov 3-4), STS-102 (Nov 5-6), STS-103 (Nov 7-8), STS-104 (Nov 9-10), STS-105 (Nov 11-12), STS-106 (Nov 13-14), STS-107 (Nov 15-16), STS-108 (Nov 17-18), STS-109 (Nov 19-20), STS-110 (Nov 21-22), STS-111 (Nov 23-24), STS-112 (Nov 25-26), STS-113 (Nov 27-28), STS-114 (Nov 29-30), STS-115 (Dec 1-2), STS-116 (Dec 3-4), STS-117 (Dec 5-6), STS-118 (Dec 7-8), STS-119 (Dec 9-10), STS-120 (Dec 11-12), STS-121 (Dec 13-14), STS-122 (Dec 15-16), STS-123 (Dec 17-18), STS-124 (Dec 19-20), STS-125 (Dec 21-22), STS-126 (Dec 23-24), STS-127 (Dec 25-26), STS-128 (Dec 27-28), STS-129 (Dec 29-30), STS-130 (Jan 1-2), STS-131 (Jan 3-4), STS-132 (Jan 5-6), STS-133 (Jan 7-8), STS-134 (Jan 9-10), STS-135 (Jan 11-12), STS-136 (Jan 13-14), STS-137 (Jan 15-16), STS-138 (Jan 17-18), STS-139 (Jan 19-20), STS-140 (Jan 21-22), STS-141 (Jan 23-24), STS-142 (Jan 25-26), STS-143 (Jan 27-28), STS-144 (Jan 29-30), STS-145 (Feb 1-2), STS-146 (Feb 3-4), STS-147 (Feb 5-6), STS-148 (Feb 7-8), STS-149 (Feb 9-10), STS-150 (Feb 11-12), STS-151 (Feb 13-14), STS-152 (Feb 15-16), STS-153 (Feb 17-18), STS-154 (Feb 19-20), STS-155 (Feb 21-22), STS-156 (Feb 23-24), STS-157 (Feb 25-26), STS-158 (Feb 27-28), STS-159 (Feb 29-Mar 1), STS-160 (Mar 3-4), STS-161 (Mar 5-6), STS-162 (Mar 7-8), STS-163 (Mar 9-10), STS-164 (Mar 11-12), STS-165 (Mar 13-14), STS-166 (Mar 15-16), STS-167 (Mar 17-18), STS-168 (Mar 19-20), STS-169 (Mar 21-22), STS-170 (Mar 23-24), STS-171 (Mar 25-26), STS-172 (Mar 27-28), STS-173 (Mar 29-30), STS-174 (Apr 1-2), STS-175 (Apr 3-4), STS-176 (Apr 5-6), STS-177 (Apr 7-8), STS-178 (Apr 9-10), STS-179 (Apr 11-12), STS-180 (Apr 13-14), STS-181 (Apr 15-16), STS-182 (Apr 17-18), STS-183 (Apr 19-20), STS-184 (Apr 21-22), STS-185 (Apr 23-24), STS-186 (Apr 25-26), STS-187 (Apr 27-28), STS-188 (Apr 29-30), STS-189 (May 1-2), STS-190 (May 3-4), STS-191 (May 5-6), STS-192 (May 7-8), STS-193 (May 9-10), STS-194 (May 11-12), STS-195 (May 13-14), STS-196 (May 15-16), STS-197 (May 17-18), STS-198 (May 19-20

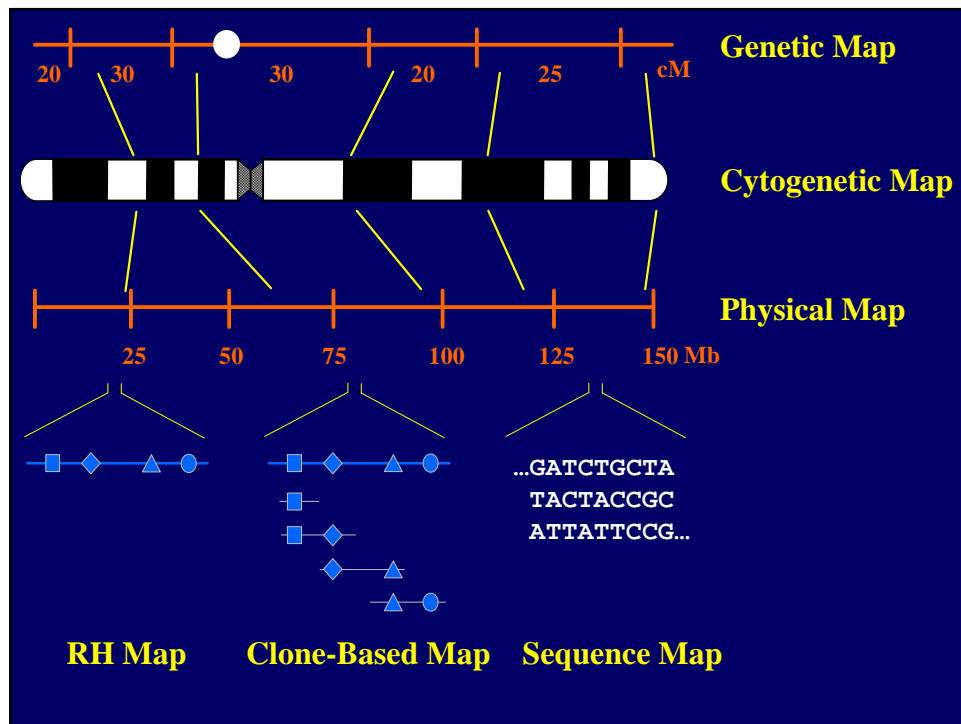
15

STSs as Landmarks

- Advantages of STSs as Landmarks
 - Independent of the Mapping Resource (Clones, RH Panel)
 - PCR-based (Sensitivity, Specificity, Automation)
 - Electronic-based Nature of STSs
 - Sequence-based Nature Facilitates Integration with Sequence
- General Review on STS-Content Mapping: Green and Green (1991)
- Programmatic Goal of U.S. Human Genome Project [Collins & Galas (1993)]
 - 100-kb Average Resolution STS Map of Human Genome
 - Therefore, ~30,000 STSs for Human Genome
- STS Map as 'Intermediate Map' En Route to Sequencing
- Conceptual Similarity of STSs and Probes

Development of STSs

- Operational Definition of an STS
 1. Sequence that Can be Amplified by a PCR Assay
 2. Functionally is 'Unique' in the Genome
- DNA Sequence → Select Primers → Confirm Above Definition
- Generation of Sequence for Developing STSs (see Vollrath 1999)
 1. Non-Targeted (i.e., Genome-Wide)
 2. Targeted
- Targeted Approaches
 - Specific Chromosomes
 - Somatic Hybrid Cell Lines
 - Flow Sorting
 - Microdissection
 - Genetic Markers (Microsatellites)
 - Expressed Sequences [Genes, ESTs]



Map Integration

· Rationale:

Maximizes Utility of Maps

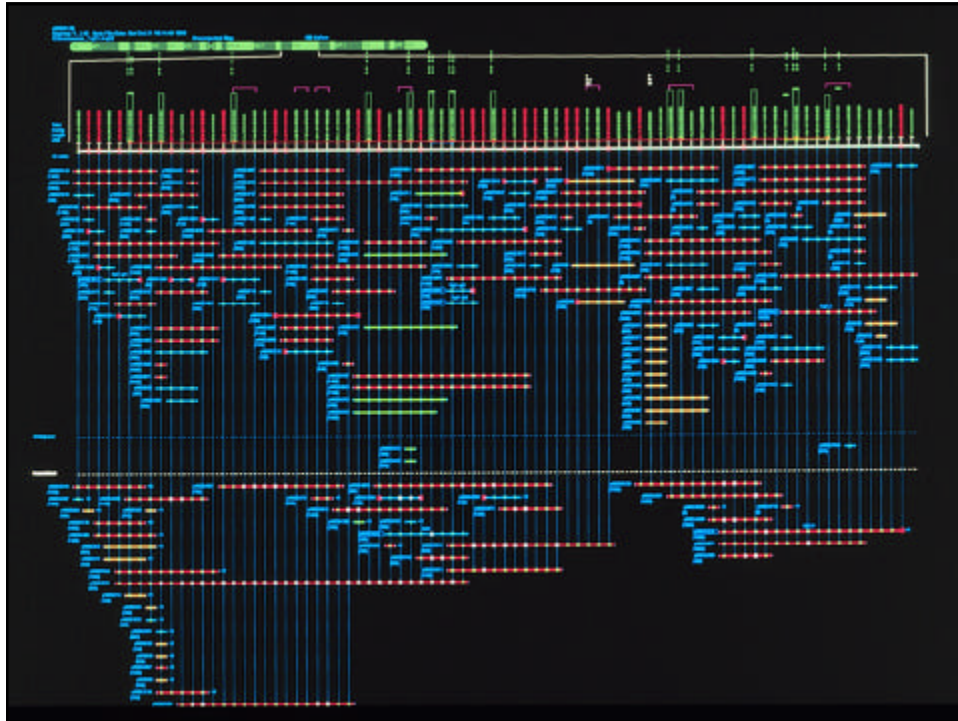
Assists in the Construction of Maps
(i.e., Provides 'Cross-Checks')

· Types of Map Integration:

Physical & Genetic

Physical & Cytogenetic

Genetic & Cytogenetic



1st Generation Clone-based Physical Maps of the Human Genome

- Constructed with YACs
- Genome-Wide Efforts

**CEPH-Genethon
Whitehead/MIT**

- Chromosome-Specific Efforts

CEPH/Genethon YAC Map of Human Genome

- Bellanne-Chantelot et al. (1992), Cohen et al. (1993), Chumakov et al. (1995)

- Experimental Data Set

- Hybridization-based Fingerprints

- Hybridization Analysis (YAC x YAC) via *Alu*-PCR

- Alu*-PCR Hybridization Assignment of YACs to Chromosomes

- FISH-Based Assignment of YACs to Chromosomes

- ***Assignment of Genethon Genetic Markers (STSs) to YACs

- Data Analysis

- Complicated!!!

- Suite of Programs to 'Disambiguate' the Data (*Quickmap*)

- Heavy Reliance on Genethon Genetic Map for Contig Assembly

- Predict 'Most Likely' Paths Among Overlapping Clones

- Map Highlights

- 225 Contigs Averaging 10 Mb, ~75% of Genome Covered

- Potentially Useful for Positional Cloning Projects

- Poor Scaffold for DNA Sequencing (Sparse STS Density)

- Data and Map Availability: www.cephb.fr/bio/ceph-genethon-map.html

Whitehead/MIT YAC Map of Human Genome

- Hudson et al. (1995)

- ~25,000 STSs Mapped Relative to YACs and/or RH Panel and/or by Genetic Mapping

- Integrated Approach for Physical Mapping of STSs

- 1. YAC-based STS-Content Mapping

- ~11,000 STSs, CEPH Mega-YACs

- 2. RH Mapping of STSs

- ~15,000 STSs, GeneBridge 4 RH Panel

- 3. Genethon Genetic Maps

- ~5,300 STSs

- PCR Analysis: *Genomatron*

- Massively-Parallel, Factory-Style Automation System

- 1,536 Position Arrays

- ~150,000 PCR Assays Per Run

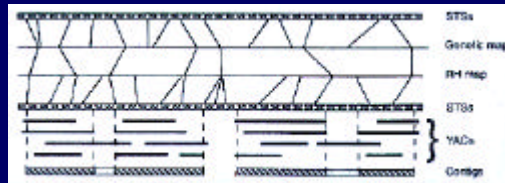
- >25,000,000 PCR Assays Total

Whitehead/MIT YAC Map of Human Genome

- Strategy for Map Construction

Genetic and RH Maps Provide Global Framework ('Top-Down Mapping')

YAC-based STS-Content Map Provides Local Ordering of STSs ('Bottom-Up Mapping')



Cross-Reference to Deduce an 'Integrated Map' of Each Chromosome

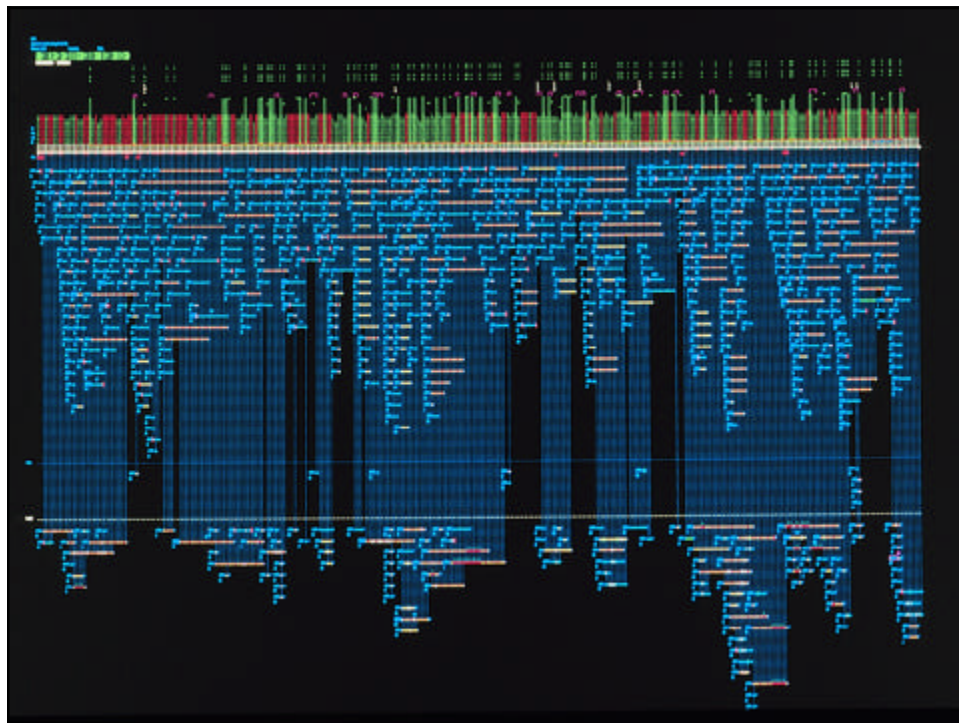
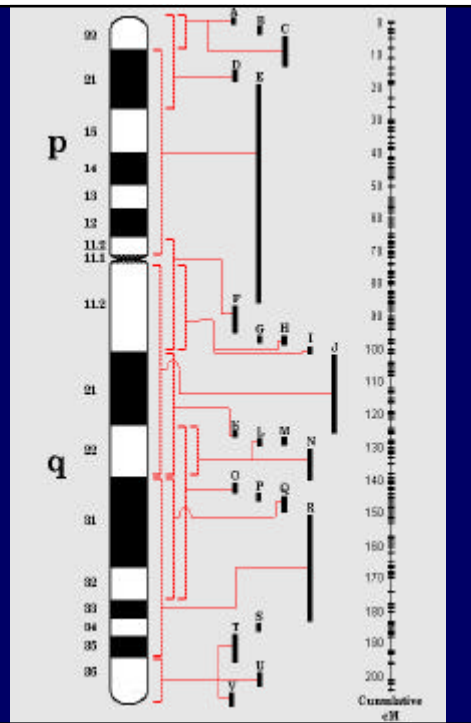
- Average STS Resolution: ~120 kb
- Availability of Data and Maps: www-genome.wi.mit.edu

Chromosome-Specific YAC Maps of Human Genome

- Chromosome 3 Gemmill et al. (1995)
- Chromosome 4 Goold et al. (1993)
- Chromosome 7 Green et al. (1994, 1995), Bouffard et al. (1997)
- Chromosome 10 Genome Therapeutics, Unpublished
- Chromosome 11 Smith et al. (1993), Quackenbush et al. (1995), Qin et al. (1996)
- Chromosome 12 Krauter et al. (1995)
- Chromosome 16 Doggett et al. (1995)
- Chromosome 19 Ashworth et al. (1995)
- Chromosome 21 Chumakov et al. (1992), Korenberg et al. (1995), Wang et al. (1999)
- Chromosome 22 Bell et al. (1995), Collins et al. (1995)
- Chromosome X Nagaraja et al. (1997)
- Chromosome Y Foote et al. (1992), Vollrath et al. (1992)

YAC-based Physical Map of Human Chromosome 7

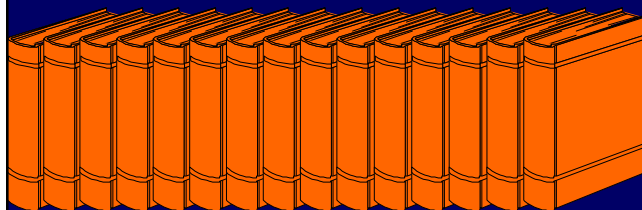
Bouffard et al. (1997)



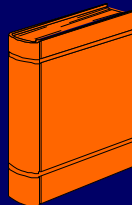
Guide to Web Sites with Physical Mapping Data



www.nhgri.nih.gov/Data



Genome
(~3000 Mb)



Chromosome
(~130 Mb)

```
GATCCTCTAGAATCTC
GAGACTCTTGAGATCT
GTGGAACTCTGTGA
TCTGACTAGCCAGT

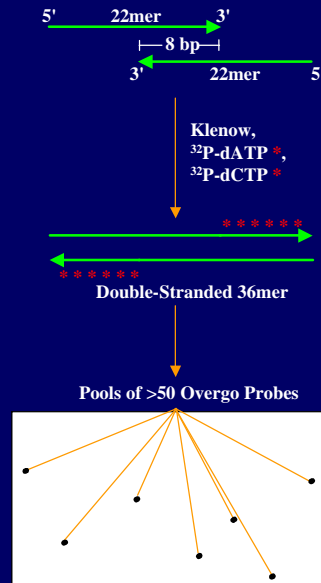
TACGTGTGAGAGATGT
ATGATGACCTGACCC
GGGTTCACTCTGAC
GACTGACCTGACCTCA

GAGGCCACCCGCGCT
GTTCACCTGACCTGAC
```

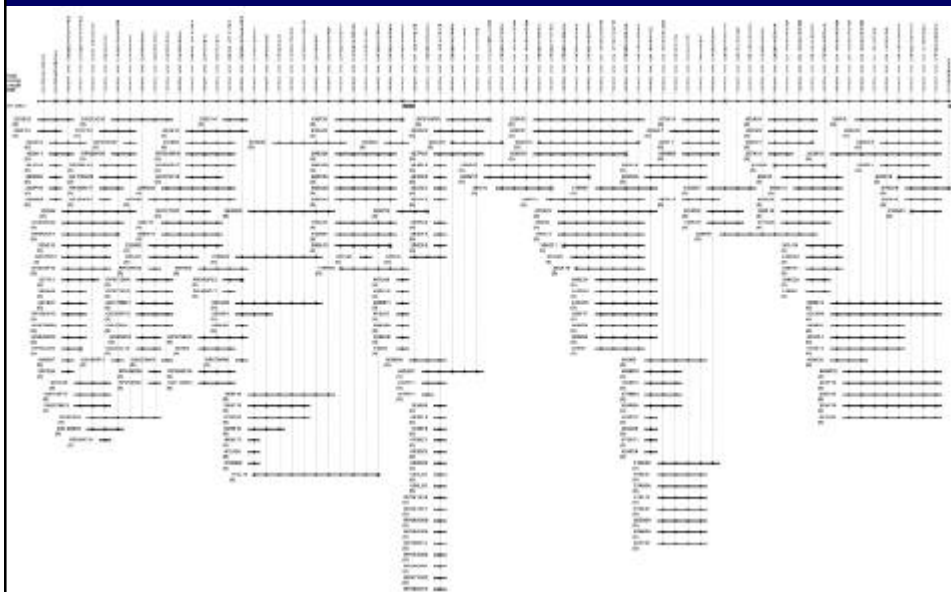
BAC
(~0.1-0.2 Mb)

'Overgo' Hybridization Probes

- Pair of ~22mer Oligonucleotide Primers with 8-bp Overlap
- Primer Extension with Klenow and both ^{32}P -dATP and ^{32}P -dCTP
- Low Background Allows Pooling of Multiple Overgo Probes



Probe-Content BAC Contig Map



Restriction Enzyme Digest-based Fingerprint Analysis

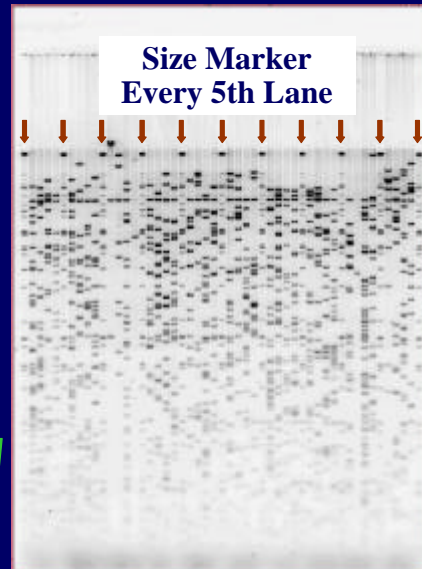
- BAC DNA Purification in 96-Well Format

- HindIII Digestion

- 1% Agarose Gel

>20 kb

~300 bp



Marra et al., 1997

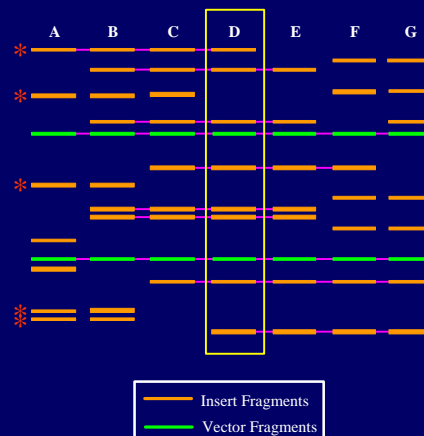
Restriction Enzyme Digest-based Fingerprint Analysis

Identify Overlapping Clones by the Presence of Common Restriction Fragments

Process Repeated Iteratively to Assemble a BAC Contig

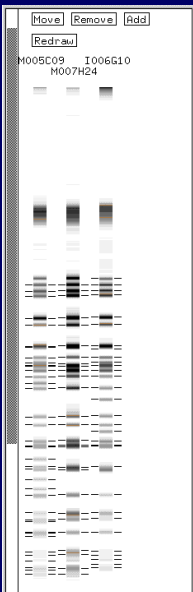
A Clone Selected for Sequencing Must Have All Restriction Fragments Accounted for in Overlapping Clones

Restriction Enzyme-Digested BAC DNA



Marra et al., 1997

IMAGE



FPC

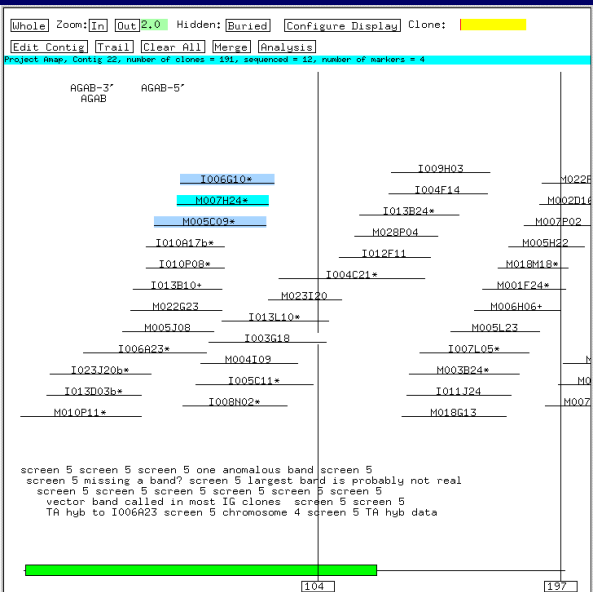
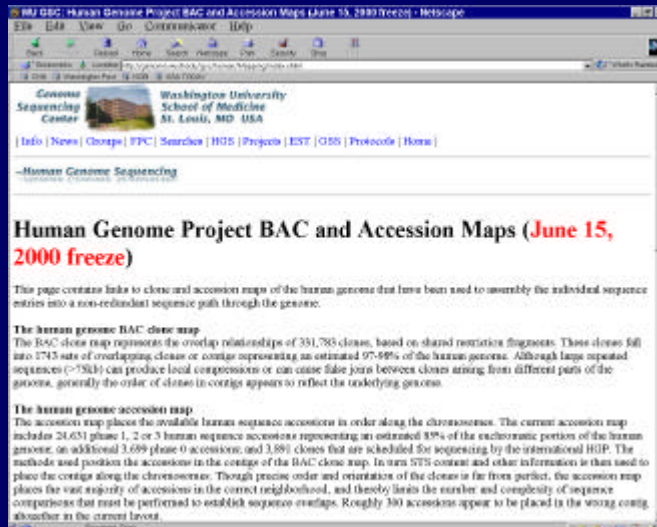


IMAGE and FPC: www.sanger.ac.uk/Software

BAC-based Physical Map of Human Genome



Human Genome Project BAC and Accession Maps (June 15, 2000 freeze)

This page contains links to clone and accession maps of the human genome that have been used to assemble the individual sequence entries into a non-redundant sequence path through the genome.

The human genome BAC clone map
 The BAC clone map represents the overlap relationships of 351,785 clones, based on shared restriction fragments. These clones fall into 1745 sets of overlapping clones or contigs representing an estimated 97-99% of the human genome. Although large repeated sequences (>15kb) can produce local compression or can cause false joins between clones arising from different parts of the genome, generally the order of clones in contigs appears to reflect the underlying genome.

The human genome accession map
 The accession map places the available human sequence accessions in order along the chromosomes. The current accession map includes 24,631 phase 1, 2 or 3 human sequence accessions representing an estimated 89% of the euchromatic portion of the human genome, an additional 3,699 phase 0 accessions, and 3,881 clones that are scheduled for sequencing by the international HGP. The methods used position the accessions in the context of the BAC clone map. In turn STS content and other information is thus used to place the contigs along the chromosomes. Though precise order and orientation of the clones is far from perfect, the accession map places the vast majority of accessions in the correct neighborhood, and thereby links the number and complexity of sequence comparisons that must be performed to establish sequence overlaps. Roughly 300 accessions appear to be placed in the wrong contig, elsewhere in the current layout.

genome.wustl.edu/gsc/human/Mapping/index.shtml

BAC-based Physical Map of Mouse Genome



www.bcgsc.bc.ca/projects/mouse_mapping

Future Prospects: Physical Mapping of Other Vertebrates

Rapidly changing strategies with increases in sequencing capabilities...

Mouse

www.bcgsc.bc.ca/projects/mouse_mapping

www.informatics.jax.org

www-genome.wi.mit.edu

Rat

www.informatics.jax.org/rat

Zebrafish

zfin.org/ZFIN

Others (non-human primates, dog, cat, cow, pig, etc.)???